STATISTICS AND PROBABILITY
7 TH GRADE

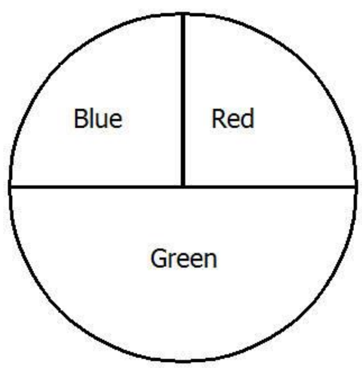
Lesson 1: Chance Experiments

Classwork

Have you ever heard a weatherman say there is a 40% chance of rain tomorrow or a football referee tell a team there is a $50/50$ chance of getting a heads on a coin toss to determine which team starts the game? These are probability statements. In this lesson, you are going to investigate probability and how likely it is that some events will occur.

Example 1: Spinner Game

Suppose you and your friend are about to play a game using the spinner shown here:



Rules of the game:

1. Decide who will go first.
2. Each person picks a color. Both players cannot pick the same color.
3. Each person takes a turn spinning the spinner and recording what color the spinner stops on. The winner is the person whose color is the first to happen 10 times.

Play the game, and remember to record the color the spinner stops on for each spin.

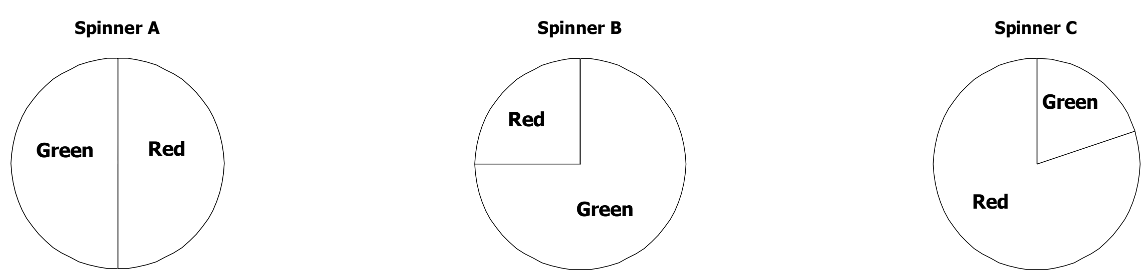
Exercises 1–4

1. Which color was the first to occur 10 times?

2. Do you think it makes a difference who goes first to pick a color?

3. Which color would you pick to give you the best chance of winning the game? Why would you pick that color?

4. Below are three different spinners. On which spinner is the green likely to win, unlikely to win, and equally likely to win?



Example 2: What is Probability?

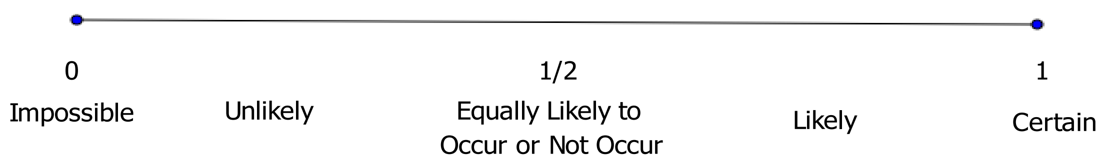
Probability is a measure of how likely it is that an event will happen. A probability is indicated by a number between 0 and 1. Some events are certain to happen, while others are impossible. In most cases, the probability of an event happening is somewhere between certain and impossible.

For example, consider a box that contains only red cubes. If you were to select one cube from the bag, you are certain to pick a red one. We say that an event that is certain to happen has a probability of 1. If we were to reach into the same bag of cubes, it is impossible to select a yellow cube. An impossible event has a probability of 0.

Description	Example	Explanation
Some events are impossible. These events have a probability of 0.	You have a bag with two green cubes, and you select one at random. Selecting a blue cube is an impossible event.	There is no way to select a blue cube if there are no blue cubes in the bag.
Some events are certain. These events have a probability of 1.	You have a bag with two green cubes, and you select one at random. Selecting a green cube is a certain event.	You will always get a green cube if there are only green cubes in the bag.
Some events are classified as equally likely to happen or to not happen. These events have a probability of $\frac{1}{2}$.	You have a bag with one blue cube and one red cube, and you randomly pick one. Selecting a blue cube is equally likely to happen or not to happen.	Since exactly half of the bag is made up of blue cubes and exactly half of the bag is comprised of red cubes, there is a 50/50 chance (equally likely) of selecting a blue cube and a 50/50 chance (equally likely) of NOT selecting a blue cube.
Some events are more likely to happen than not to happen. These events have a probability that is greater than 0.5. These events could be described as *likely* to occur.	If you have a bag that contains eight blue cubes and two red cubes, and you select one at random, it is likely that you will get a blue cube.	Even though it is not certain that you will get a blue cube, a blue cube would be selected most of the time because there are many more blue cubes than red cubes.
Some events are less likely to happen than not to happen. These events have a probability that is less than 0.5. These events could be described as *unlikely* to occur.	If you have a bag that contains eight blue cubes and two red cubes, and you select one at random, it is unlikely that you will get a red cube.	Even though it is not impossible to get a red cube, a red cube would not be selected very often because there are many more blue cubes than red cubes.

The figure below shows the probability scale.

Probability Scale



Exercises 5–10

5. Decide where each event would be located on the scale above. Place the letter for each event on the appropriate place on the probability scale.

   Event:

   A. You will see a live dinosaur on the way home from school today.

   B. A solid rock dropped in the water will sink.

   C. A round disk with one side red and the other side yellow will land yellow side up when flipped.

   D. A spinner with four equal parts numbered 1–4 will land on the 4 on the next spin.

   E. Your full name will be drawn when a full name is selected randomly from a bag containing the full names of all of the students in your class.

   F. A red cube will be drawn when a cube is selected from a bag that has five blue cubes and five red cubes.

   G. Tomorrow the temperature outside will be -250 degrees.

6. Design a spinner so that the probability of spinning green is 1.

7. Design a spinner so that the probability of spinning green is 0.

8. Design a spinner with two outcomes in which it is equally likely to land on the red and green parts.

An event that is impossible has a probability of 0 and will never occur, no matter how many observations you make. This means that in a long sequence of observations, it will occur 0% of the time. An event that is certain has a probability of 1 and will always occur. This means that in a long sequence of observations, it will occur 100% of the time.

9. What do you think it means for an event to have a probability of $\frac{1}{2}$?

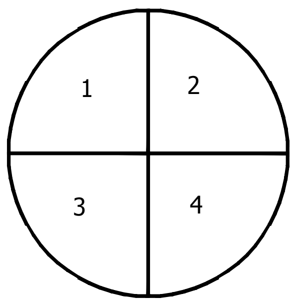
10. What do you think it means for an event to have a probability of $\frac{1}{4}$?

Lesson 2: Estimating Probabilities by Collecting Data

Classwork

Example 1: Carnival Game

At the school carnival, there is a game in which students spin a large spinner. The spinner has four equal sections numbered 1–4 as shown below. To play the game, a student spins the spinner twice and adds the two numbers that the spinner lands on. If the sum is greater than or equal to 5, the student wins a prize.



Exercises 1–8

You and your partner will play this game 15 times. Record the outcome of each spin in the table below.

Turn	1st Spin Results	2nd Spin Results	Sum
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			
13			
14			
15			

1. Out of the 15 turns how many times was the sum greater than or equal to 5?

2. What sum occurred most often?

3. What sum occurred least often?

4. If students were to play a lot of games, what proportion of the games would they win? Explain your answer.

5. Name a sum that would be impossible to get while playing the game.

6. What event is certain to occur while playing the game?

When you were spinning the spinner and recording the outcomes, you were performing a *chance experiment*. You can use the results from a chance experiment to estimate the probability of an event. In the example above, you spun the spinner 15 times and counted how many times the sum was greater than or equal to 5. An estimate for the probability of a sum greater than or equal to 5 is

$$P(\text{sum } \geq 5) = \frac{\text{Number of observed occurrences of the event}}{\text{Total number of observations}}.$$

7. Based on your experiment of playing the game, what is your estimate for the probability of getting a sum of 5 or more?

8. Based on your experiment of playing the game, what is your estimate for the probability of getting a sum of exactly 5?

Example 2: Animal Crackers

A student brought a very large jar of animal crackers to share with students in class. Rather than count and sort all the different types of crackers, the student randomly chose 20 crackers and found the following counts for the different types of animal crackers. Estimate the probability of selecting a zebra.

Lion	2
Camel	1
Monkey	4
Elephant	5
Zebra	3
Penguin	3
Tortoise	2
	Total 20

Exercises 9–15

If a student randomly selected a cracker from a large jar:

9. What is your estimate for the probability of selecting a lion?

10. What is your estimate for the probability of selecting a monkey?

11. What is your estimate for the probability of selecting a penguin or a camel?

12. What is your estimate for the probability of selecting a rabbit?

13. Is there the same number of each kind of animal cracker in the large jar? Explain your answer.

14. If the student randomly selected another 20 animal crackers, would the same results occur? Why or why not?

15. If there are 500 animal crackers in the jar, how many elephants are in the jar? Explain your answer.

Lesson 3: Chance Experiments with Equally Likely Outcomes

Classwork

Example 1

Jamal, a seventh grader, wants to design a game that involves tossing paper cups. Jamal tosses a paper cup five times and records the outcome of each toss. An ***outcome*** is the result of a single trial of an experiment.

Here are the results of each toss:

Jamal noted that the paper cup could land in one of three ways: on its side, right side up, or upside down. The collection of these three outcomes is called the *sample space* of the experiment. The ***sample space*** of an experiment is the set of all possible outcomes of that experiment.

For example, the sample space when flipping a coin is heads, tails.

The sample space when drawing a colored cube from a bag that has 3 red, 2 blue, 1 yellow, and 4 green cubes is red, blue, yellow, green.

Exercises 1–6

For each of the following chance experiments, list the sample space (i.e., all the possible outcomes).

1. Drawing a colored cube from a bag with 2 green, 1 red, 10 blue, and 3 black.

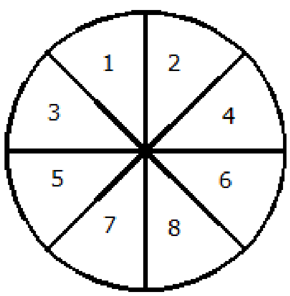
2. Tossing an empty soup can to see how it lands.

3. Shooting a free-throw in a basketball game.

4. Rolling a number cube with the numbers 1–6 on its faces.

5. Selecting a letter from the word *probability*.

6. Spinning the spinner:



Example 2: Equally Likely Outcomes

The sample space for the paper cup toss was on its side, right side up, and upside down. Do you think each of these outcomes has the same chance of occurring? If they do, then they are equally likely to occur.

The outcomes of an experiment are equally likely to occur when the probability of each outcome is equal.

Toss the paper cup 30 times and record in a table the results of each toss.

Toss	Outcome
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	
16	
17	
18	
19	
20	
21	
22	
23	
24	
25	
26	
27	
28	
29	
30	

Exercises 7–10

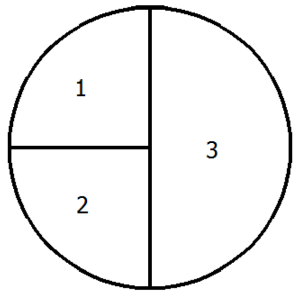
7. Using the results of your experiment, what is your estimate for the probability of a paper cup landing on its side?

8. Using the results of your experiment, what is your estimate for the probability of a paper cup landing upside down?

9. Using the results of your experiment, what is your estimate for the probability of a paper cup landing right side up?

10. Based on your results, do you think the three outcomes are equally likely to occur?

11. Using the spinner below, answer the following questions.



a. Are the events spinning and landing on 1 or 2 equally likely?

b. Are the events spinning and landing on 2 or 3 equally likely?

c. How many times do you predict the spinner will land on each section after 100 spins?

12. Draw a spinner that has 3 sections that are equally likely to occur when the spinner is spun. How many times do you think the spinner will land on each section after 100 spins?

Lesson 4: Calculating Probabilities for Chance Experiments with Equally Likely Outcomes

Classwork

Examples: Theoretical Probability

In a previous lesson, you saw that to find an estimate of the probability of an event for a chance experiment you divide

$$P(\text{event}) = \frac{\text{Number of observed occurrences of the event}}{\text{Total number of observations}}.$$

Your teacher has a bag with some cubes colored yellow, green, blue, and red. The cubes are identical except for their color. Your teacher will conduct a chance experiment by randomly drawing a cube with replacement from the bag. Record the outcome of each draw in the table below.

Trial	Outcome
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	
16	
17	
18	
19	
20	

1. Based on the 20 trials, estimate for the probability of
   a. choosing a yellow cube.

   b. choosing a green cube.

   c. choosing a red cube.

   d. choosing a blue cube.

2. If there are 40 cubes in the bag, how many cubes of each color are in the bag? Explain.

3. If your teacher were to randomly draw another 20 cubes one at a time and with replacement from the bag, would you see exactly the same results? Explain.

4. Find the fraction of each color of cubes in the bag.

   Yellow

   Green

   Red

   Blue

Each fraction is the *theoretical probability* of choosing a particular color of cube when a cube is randomly drawn from the bag.

When all the possible outcomes of an experiment are equally likely, the probability of each outcome is

$$P(\text{outcome}) = \frac{1}{\text{Number of possible outcomes}}.$$

An event is a collection of outcomes, and when the outcomes are equally likely, the theoretical probability of an event can be expressed as

$$P(\text{event}) = \frac{\text{Number of favorable outcomes}}{\text{Number of possible outcomes}}.$$

The theoretical probability of drawing a blue cube is

$$P(\text{blue}) = \frac{\text{Number of blue cubes}}{\text{Total number of cubes}} = \frac{10}{40}.$$

5. Is each color equally likely to be chosen? Explain your answer.

6. How do the theoretical probabilities of choosing each color from Exercise 4 compare to the experimental probabilities you found in Exercise 1?

7. An experiment consisted of flipping a nickel and a dime. The first step in finding the theoretical probability of obtaining a heads on the nickel and a heads on the dime is to list the sample space. For this experiment, the sample space is shown below.

Nickel	Dime
H	H
H	T
T	H
T	T

If the counts are fair, these outcomes are equally likely, so the probability of each outcome is $\frac{1}{4}$.

Nickel	Dime	Probability
H	H	$\frac{1}{4}$
H	T	$\frac{1}{4}$
T	H	$\frac{1}{4}$
T	T	$\frac{1}{4}$

The probability of two heads is $\frac{1}{4}$ or $P(\text{two heads}) = \frac{1}{4}$.

Exercises

8. Consider a chance experiment of rolling a number cube.

   a. What is the sample space? List the probability of each outcome in the sample space.

   b. What is the probability of rolling an odd number?

   c. What is the probability of rolling a number less than 5?

9. Consider an experiment of randomly selecting a letter from the word *number*.

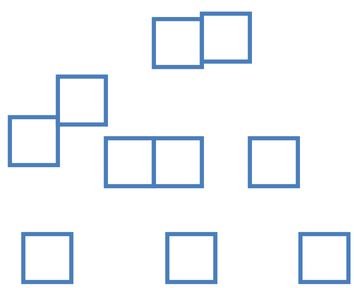
 a. What is the sample space? List the probability of each outcome in the sample space.

 b. What is the probability of selecting a vowel?

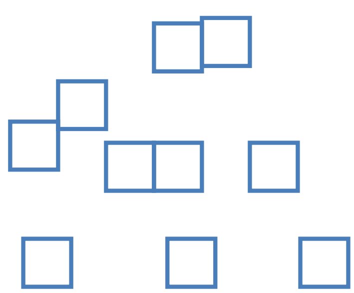
 c. What is the probability of selecting the letter z?

10. Consider an experiment of randomly selecting a cube from a bag of 10 cubes.

 a. Color the cubes below so that the probability of selecting a blue cube is $\frac{1}{2}$.



 b. Color the cubes below so that the probability of selecting a blue cube is $\frac{4}{5}$.



11. Students are playing a game that requires spinning the two spinners shown below. A student wins the game if both spins land on red. What is the probability of winning the game? Remember to first list the sample space and the probability of each outcome in the sample space. There are eight possible outcomes to this chance experiment.

Lesson 5: Chance Experiments with Outcomes That Are Not Equally Likely

Classwork

In previous lessons, you learned that when the outcomes in a sample space are equally likely, the probability of an event is the number of outcomes in the event divided by the number of outcomes in the sample space. However, when the outcomes in the sample space are *not* equally likely, we need to take a different approach.

Example 1

When Jenna goes to the farmer's market she usually buys bananas. The number of bananas she might buy and their probabilities are shown in the table below.

Number of Bananas	0	1	2	3	4	5
Probability	0.1	0.1	0.1	0.2	0.2	0.3

a. What is the probability that Jenna buys exactly 3 bananas?

b. What is the probability that Jenna doesn't buy any bananas?

c. What is the probability that Jenna buys more than 3 bananas?

d. What is the probability that Jenna buys at least 3 bananas?

e. What is the probability that Jenna doesn't buy exactly 3 bananas?

Notice that the sum of the probabilities in the table is one whole $(0.1 + 0.1 + 0.1 + 0.2 + 0.2 + 0.3 = 1)$. This is always true; when we add up the probabilities of all the possible outcomes, the result is always 1. So, taking 1 and subtracting the probability of the event gives us the probability of something NOT occurring.

Exercises 1–2

Jenna's husband, Rick, is concerned about his diet. On any given day, he eats 0, 1, 2, 3, or 4 servings of fruit and vegetables. The probabilities are given in the table below.

Number of Servings of Fruit and Vegetables	0	1	2	3	4
Probability	0.08	0.13	0.28	0.39	0.12

1. On a given day, find the probability that Rick eats:
   a. Two servings of fruit and vegetables.

   b. More than two servings of fruit and vegetables.

   c. At least two servings of fruit and vegetables.

2. Find the probability that Rick does not eat exactly two servings of fruit and vegetables.

Example 2

Luis works in an office, and the phone rings occasionally. The possible number of phone calls he receives in an afternoon and their probabilities are given in the table below.

Number of Phone Calls	0	1	2	3	4
Probability	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{2}{9}$	$\frac{1}{3}$	$\frac{1}{9}$

a. Find the probability that Luis receives 3 or 4 phone calls.

b. Find the probability that Luis receives fewer than 2 phone calls.

c. Find the probability that Luis receives 2 or fewer phone calls.

d. Find the probability that Luis does not receive 4 phone calls.

Exercises 3–7

When Jenna goes to the farmer's market, she also usually buys some broccoli. The possible number of heads of broccoli that she buys and the probabilities are given in the table below.

Number of Heads of Broccoli	0	1	2	3	4
Probability	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{5}{12}$	$\frac{1}{4}$	$\frac{1}{12}$

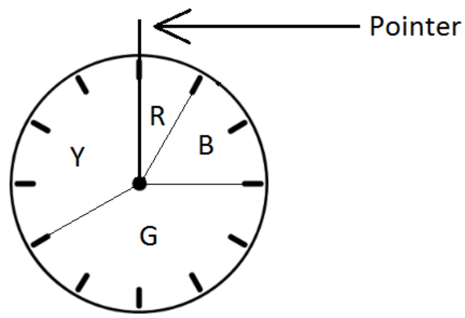
3. Find the probability that Jenna
   a. Buys exactly 3 heads of broccoli.

   b. Does not buy exactly 3 heads of broccoli.

   c. Buys more than 1 head of broccoli.

   d. Buys at least 3 heads of broccoli.

The diagram below shows a spinner designed like the face of a clock. The sectors of the spinner are colored red (R), blue (B), green (G), and yellow (Y).



4. Writing your answers as fractions in lowest terms, find the probability that the pointer stops on the following colors.
   a. Red:

   b. Blue:

   c. Green:

   d. Yellow:

5. Complete the table of probabilities below.

Color	Red	Blue	Green	Yellow
Probability				

6. Find the probability that the pointer stops in either the blue region or the green region.

7. Find the probability that the pointer does not stop in the green region.

Lesson 6: Using Tree Diagrams to Represent a Sample Space and to Calculate Probabilities

Classwork

Suppose a girl attends a preschool where the students are studying primary colors. To help teach calendar skills, the teacher has each student maintain a calendar in his or her cubby. For each of the four days that they are covering primary colors in class, students get to place a colored dot on their calendar: blue, yellow, or red. When the four days of the school week have passed (Monday–Thursday), what might the young girl's calendar look like?

One outcome would be four blue dots if the student chose blue each day. But consider that the first day (Monday) could be blue, and the next day (Tuesday) could be yellow, and Wednesday could be blue, and Thursday could be red. Or, maybe Monday and Tuesday could be yellow, Wednesday could be blue, and Thursday could be red. Or, maybe Monday, Tuesday, and Wednesday could be blue, and Thursday could be red, and so on and so forth.

As hard to follow as this seems now, we have only mentioned 3 of the 81 possible outcomes in terms of the four days of colors! Listing the other 78 outcomes would take several pages! Rather than listing outcomes in the manner described above (particularly when the situation has multiple stages, such as the multiple days in the case above), we often use a *tree diagram* to display all possible outcomes visually. Additionally, when the outcomes of each stage are the result of a chance experiment, tree diagrams are helpful for computing probabilities.

Example 1: Two Nights of Games

Imagine that a family decides to play a game each night. They all agree to use a tetrahedral die (i.e., a four-sided pyramidal die where each of four possible outcomes is equally likely—see image on page S.44) each night to randomly determine if they will play a board game (B) or a card game (C). The tree diagram mapping the possible overall outcomes over two consecutive nights will be developed below.

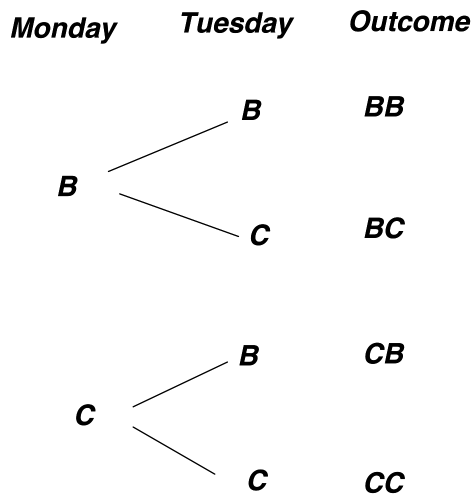
To make a tree diagram, first present all possibilities for the first stage. (In this case, Monday.)

Monday ***Tuesday*** ***Outcome***

B

C

Then, from *each* branch of the first stage, attach all possibilities for the second stage (Tuesday).



Note: If the situation has more than two stages, this process would be repeated until all stages have been presented.

a. If BB represents two straight nights of board games, what does CB represent?

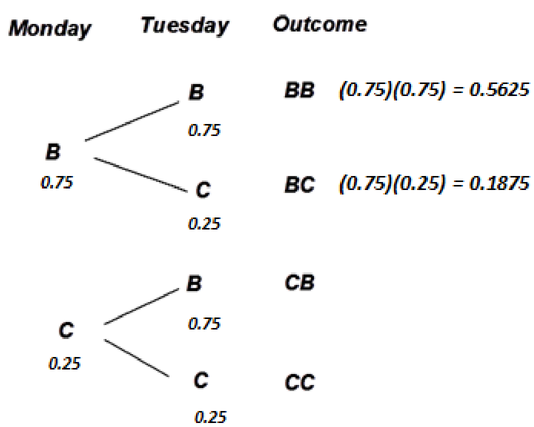
b. List the outcomes where exactly one board game is played over two days. How many outcomes were there?

Example 2: Two Nights of Games (with Probabilities)

In the example above, each night's outcome is the result of a chance experiment (rolling the tetrahedral die). Thus, there is a probability associated with each night's outcome.

By multiplying the probabilities of the outcomes from each stage, we can obtain the probability for each "branch of the tree." In this case, we can figure out the probability of each of our four outcomes: BB, BC, CB, and CC.

For this family, a card game will be played if the die lands showing a value of 1, and a board game will be played if the die lands showing a value of 2, 3, or 4. This makes the probability of a board game (B) on a given night 0.75.



a. The probabilities for two of the four outcomes are shown. Now, compute the probabilities for the two remaining outcomes.

c. What is the probability that there will be exactly one night of board games over the two nights?

Exercises: Two Children

Two friends meet at a grocery store and remark that a neighboring family just welcomed their second child. It turns out that both children in this family are girls, and they are not twins. One of the friends is curious about what the chances are of having 2 girls in a family's first 2 births. Suppose that for each birth the probability of a boy birth is 0.5 and the probability of a girl birth is also 0.5.

1. Draw a tree diagram demonstrating the four possible birth outcomes for a family with 2 children (no twins). Use the symbol B for the outcome of *boy* and G for the outcome of *girl*. Consider the first birth to be the first stage. (Refer to Example 1 if you need help getting started.)

2. Write in the probabilities of each stage's outcome to the tree diagram you developed above, and determine the probabilities for each of the 4 possible birth outcomes for a family with 2 children (no twins).

3. What is the probability of a family having 2 girls in this situation? Is that greater than or less than the probability of having exactly 1 girl in 2 births?

Lesson 7: Calculating Probabilities of Compound Events

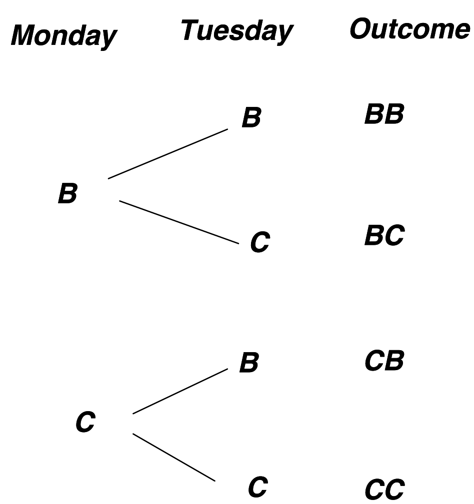
Classwork

A previous lesson introduced *tree diagrams* as an effective method of displaying the possible outcomes of certain multistage chance experiments. Additionally, in such situations, tree diagrams were shown to be helpful for computing probabilities.

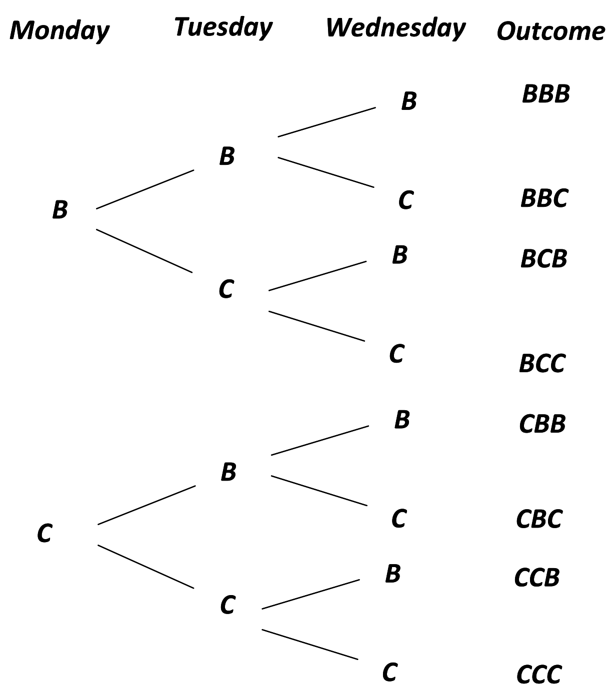
In those previous examples, diagrams primarily focused on cases with two stages. However, the basic principles of tree diagrams can apply to situations with more than two stages.

Example 1: Three Nights of Games

Recall a previous example where a family decides to play a game each night, and they all agree to use a tetrahedral die (a four-sided die in the shape of a pyramid where each of four possible outcomes is equally likely) each night to randomly determine if the game will be a board (B) or a card (C) game. The tree diagram mapping the possible overall outcomes over two consecutive nights was as follows:



But how would the diagram change if you were interested in mapping the possible overall outcomes over three consecutive nights? To accommodate this additional third stage, you would take steps similar to what you did before. You would attach all possibilities for the third stage (Wednesday) to each branch of the previous stage (Tuesday).



Exercises 1–3

1. If BBB represents three straight nights of board games, what does CBB represent?

2. List all outcomes where exactly two board games were played over three days. How many outcomes were there?

3. There are eight possible outcomes representing the three nights. Are the eight outcomes representing the three nights equally likely? Why or why not?

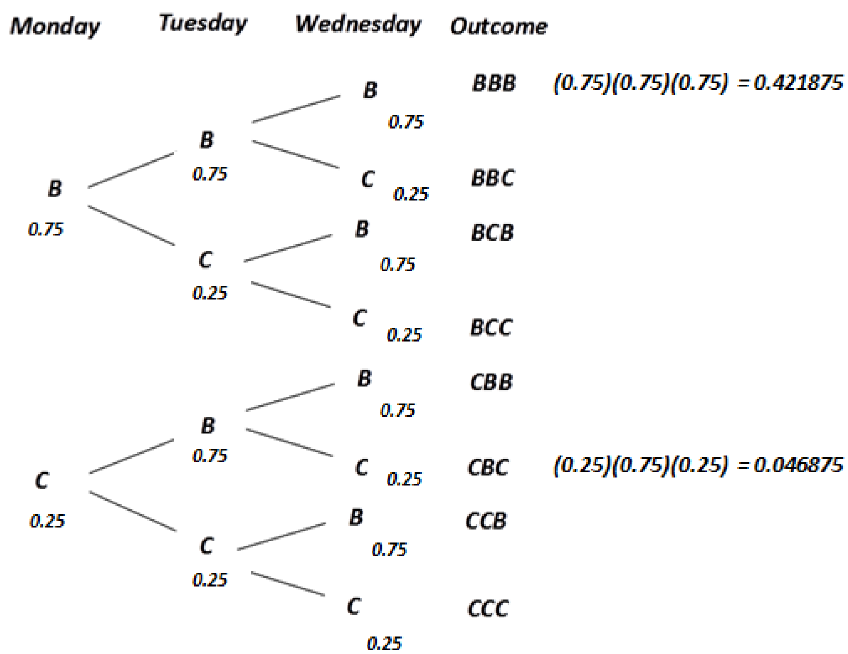
Example 2: Three Nights of Games (with Probabilities)

In the example above, each night's outcome is the result of a chance experiment (rolling the four-sided die). Thus, there is a probability associated with each night's outcome.

By multiplying the probabilities of the outcomes from each stage, you can obtain the probability for each "branch of the tree." In this case, you can figure out the probability of each of our eight outcomes.

For this family, a card game will be played if the die lands showing a value of 1, and a board game will be played if the die lands showing a value of 2, 3, or 4. This makes the probability of a board game (B) on a given night 0.75.

Let's use a tree to examine the probabilities of the outcomes for the three days.



Exercises 4–6

4. Probabilities for two of the eight outcomes are shown. Calculate the approximate probabilities for the remaining six outcomes.

5. What is the probability that there will be exactly two nights of board games over the three nights?

6. What is the probability that the family will play at least one night of card games?

Exercises 7–10: Three Children

A neighboring family just welcomed their third child. It turns out that all 3 of the children in this family are girls, and they are not twins or triplets. Suppose that for each birth the probability of a boy birth is 0.5 and the probability of a girl birth is also 0.5. What are the chances of having 3 girls in a family's first 3 births?

7. Draw a tree diagram showing the eight possible birth outcomes for a family with 3 children (no twins or triplets). Use the symbol B for the outcome of *boy* and G for the outcome of *girl*. Consider the first birth to be the first stage. (Refer to Example 1 if you need help getting started.)

8. Write in the probabilities of each stage's outcomes in the tree diagram you developed above, and determine the probabilities for each of the eight possible birth outcomes for a family with 3 children (no twins).

9. What is the probability of a family having 3 girls in this situation? Is that greater than or less than the probability of having exactly 2 girls in 3 births?

10. What is the probability of a family of 3 children having at least 1 girl?

Lesson 8: The Difference Between Theoretical Probabilities and Estimated Probabilities

Classwork

Did you ever watch the beginning of a Super Bowl game? After the traditional handshakes, a coin is tossed to determine which team gets to kick-off first. Whether or not you are a football fan, the toss of a fair coin is often used to make decisions between two groups.

Examples 1–9: Why a Coin?

Coins were discussed in previous lessons of this module. What is special about a coin? In most cases, a coin has two different sides: a head side (heads) and a tail side (tails). The sample space for tossing a coin is {heads, tails}. If each outcome has an equal chance of occurring when the coin is tossed, then the probability of getting heads is $\frac{1}{2}$, or 0.5. The probability of getting tails is also 0.5. Note that the sum of these probabilities is 1.

The probabilities formed using the sample space and what we know about coins are called the theoretical probabilities. Using observed relative frequencies is another method to estimate the probabilities of heads or tails. A relative frequency is the proportion derived from the number of the observed outcomes of an event divided by the total number of outcomes. Recall from earlier lessons that a relative frequency can be expressed as a fraction, a decimal, or a percent. Is the estimate of a probability from this method close to the theoretical probability? The following example investigates how relative frequencies can be used to estimate probabilities.

Beth tosses a coin 10 times and records her results. Here are the results from the 10 tosses:

Toss	1	2	3	4	5	6	7	8	9	10
Result	H	H	T	H	H	H	T	T	T	H

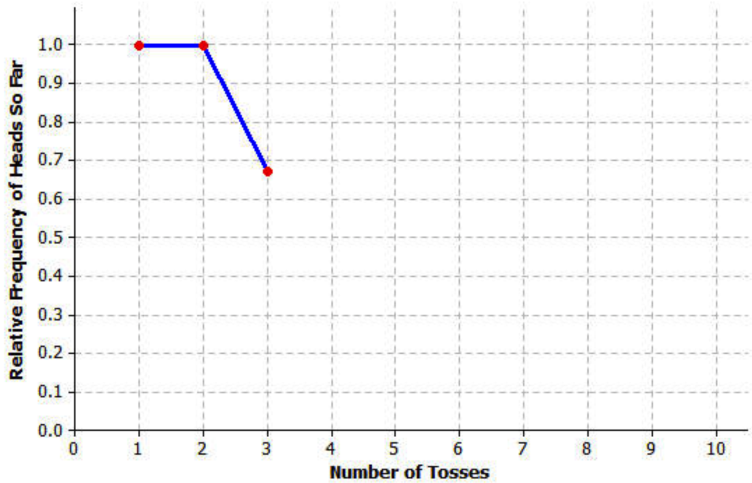
The total number of heads divided by the total number of tosses is the relative frequency of heads. It is the proportion of the time that heads occurred on these tosses. The total number of tails divided by the total number of tosses is the relative frequency of tails.

1. Beth started to complete the following table as a way to investigate the relative frequencies. For each outcome, the total number of tosses increased. The total number of heads or tails observed so far depends on the outcome of the current toss. Complete this table for the 10 tosses recorded above.

Toss	Outcome	Total number of heads so far	Relative frequency of heads so far (to the nearest hundredth)	Total number of tails so far	Relative frequency of tails so far (to the nearest hundredth)
1	H	1	$\frac{1}{1} = 1$	0	$\frac{0}{1} = 0$
2	H	2	$\frac{2}{2} = 1$	0	$\frac{0}{2} = 0$
3	T	2	$\frac{2}{3} = 0.67$	1	$\frac{1}{3} = 0.33$
4					
5					
6					
7					
8					
9					
10					

2. What is the sum of the relative frequency of heads and the relative frequency of tails for each row of the table?

3. Beth's results can also be displayed using a graph. From the table above, complete the graph below using the values of relative frequency of heads so far.



4. Beth continued tossing the coin and recording results for a total of 40 tosses. Here are the results of the next 30 tosses:

Toss	11	12	13	14	15	16	17	18	19	20
Result	T	H	T	H	T	H	H	T	H	T

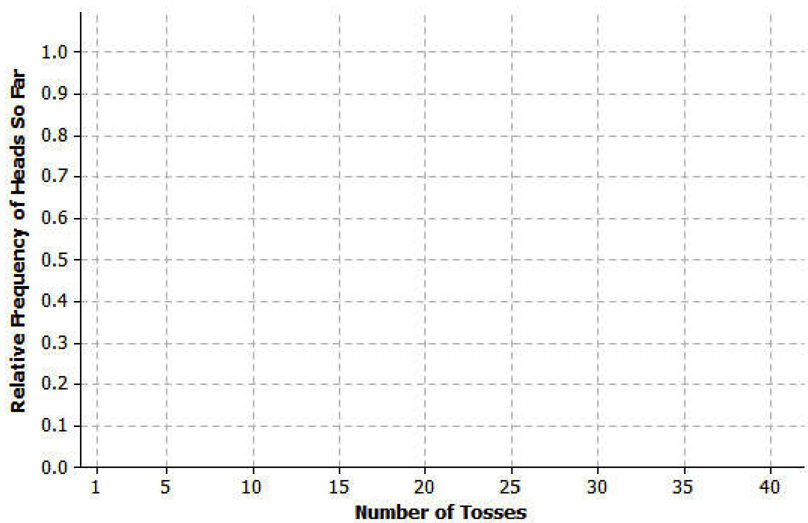
Toss	21	22	23	24	25	26	27	28	29	30
Result	H	T	T	H	T	T	T	T	H	T

Toss	31	32	33	34	35	36	37	38	39	40
Result	H	T	H	T	H	T	H	H	T	T

As the number of tosses increases, the relative frequency of heads changes. Complete the following table for the 40 coin tosses:

Number of tosses	Total number of heads so far	Relative frequency of heads so far (to the nearest hundredth)
1		
5		
10		
15		
20		
25		
30		
35		
40		

5. From the table above, complete the graph below using the relative frequency of heads so far for the total number of tosses of 1, 5, 10, 15, 20, 25, 30, 35, and 40.



6. What do you notice about the changes in the relative frequency of the number of heads so far as the number of tosses increases?

7. If you tossed the coin 100 times, what do you think the relative frequency of heads would be? Explain your answer.

8. Based on the graph and the relative frequencies, what would you estimate the probability of getting heads to be? Explain your answer.

9. How close is your estimate in Exercise 8 to the theoretical probability of 0.5? Would the estimate of this probability have been as good if Beth had only tossed the coin a few times instead of 40?

The value you gave in Exercise 8 is an estimate of the theoretical probability and is called an experimental or estimated probability.

Lesson 9: Comparing Estimated Probabilities to Probabilities Predicted by a Model

Classwork

Exploratory Challenge: Game Show—Picking Blue!

Imagine, for a moment, the following situation: You and your classmates are contestants on a quiz show called *Picking Blue!* There are two bags in front of you, Bag A and Bag B. Each bag contains red and blue chips. You are told that one of the bags has exactly the same number of blue chips as red chips. But you are told nothing about the ratio of blue to red chips in the other bag.

Each student in your class will be asked to select either Bag A or Bag B. Starting with Bag A, a chip is randomly selected from the bag. If a blue chip is drawn, all of the students in your class who selected Bag A win a Blue Token. The chip is put back in the bag. After mixing up the chips in the bag, another chip is randomly selected from the bag. If the chip is blue, the students who picked Bag A win another Blue Token. After the chip is placed back into the bag, the process continues until a red chip is picked. When a red chip is picked, the game moves to Bag B. A chip from the Bag B is then randomly selected. If it is blue, all of the students who selected Bag B win a Blue Token. But if the chip is red, the game is over. Just like for Bag A, if the chip is blue, the process repeats until a red chip is picked from the bag. When the game is over, the students with the greatest number of Blue Tokens are considered the winning team.

Without any information about the bags, you would probably select a bag simply by guessing. But surprisingly, the show's producers are going to allow you to do some research before you select a bag. For the next 20 minutes, you can pull a chip from either one of the two bags, look at the chip, and then put the chip back in the bag. You can repeat this process as many times as you want within the 20 minutes. At the end of 20 minutes, you must make your final decision and select which of the bags you want to use in the game.

Getting Started

Assume that the producers of the show do not want to give away a lot of their Blue Tokens. As a result, if one bag has the same number of red and blue chips, do you think the other bag would have more, or fewer, blue chips than red chips? Explain your answer.

Planning the Research

Your teacher will provide you with two bags labeled A and B. You have 20 minutes to experiment with pulling chips one at a time from the bags. After you examine a chip, you must put it back in the bag. Remember, no peeking in the bags as that will disqualify you from the game. You can pick chips from just one bag, or you can pick chips from one bag and then the other bag.

Use the results from 20 minutes of research to determine which bag you will choose for the game.

Provide a description outlining how you will carry out your research:

Carrying Out the Research

Share your plan with your teacher. Your teacher will verify whether your plan is within the rules of the quiz show. Approving your plan does not mean, however, that your teacher is indicating that your research method offers the most accurate way to determine which bag to select. If your teacher approves your research, carry out your plan as outlined. Record the results from your research, as directed by your teacher.

Playing the Game

After the research has been conducted, the competition begins. First, your teacher will shake up Bag A. A chip is selected. If the chip is blue, all students who selected Bag A win an imaginary Blue Token. The chip is put back in the bag, and the process continues. When a red chip is picked from Bag A, students selecting Bag A have completed the competition. Your teacher will now shake up Bag B. A chip is selected. If it is blue, all students who selected Bag B win an imaginary Blue Token. The process continues until a red chip is picked. At that point, the game is over.

How many Blue Tokens did you win?

Examining Your Results

At the end of the game, your teacher will open the bags and reveal how many blue and red chips were in each bag. Answer the questions that follow. After you have answered these questions, discuss them with your class.

1. Before you played the game, what were you trying to learn about the bags from your research?

2. What did you expect to happen when you pulled chips from the bag with the same number of blue and red chips? Did the bag that you thought had the same number of blue and red chips yield the results you expected?

3. How confident were you in predicting which bag had the same number of blue and red chips? Explain.

4. What bag did you select to use in the competition and why?

5. If you were the show's producers, how would you make up the second bag? (Remember, one bag has the same number of red and blue chips.)

6. If you picked a chip from Bag B 100 times and found that you picked each color exactly 50 times, would you know for sure that bag B was the one with equal numbers of each color?

Lesson 10: Conducting a Simulation to Estimate the Probability of an Event

Classwork

In previous lessons, you estimated probabilities of events by collecting data empirically or by establishing a theoretical probability model. There are real problems for which those methods may be difficult or not practical to use. Simulation is a procedure that will allow you to answer questions about real problems by running experiments that closely resemble the real situation.

It is often important to know the probabilities of real-life events that may not have known theoretical probabilities. Scientists, engineers, and mathematicians design simulations to answer questions that involve topics such as diseases, water flow, climate changes, or functions of an engine. Results from the simulations are used to estimate probabilities that help researchers understand problems and provide possible solutions to these problems.

Example 1: Families

How likely is it that a family with three children has all boys or all girls?

Let's assume that a child is equally likely to be a boy or a girl. Instead of observing the result of actual births, a toss of a fair coin could be used to simulate a birth. If the toss results in heads (H), then we could say a boy was born; if the toss results in tails (T), then we could say a girl was born. If the coin is fair (i.e., heads and tails are equally likely), then getting a boy or a girl is equally likely.

Exercises

Suppose that a family has three children. To simulate the genders of the three children, the coin or number cube or a card would need to be used three times, once for each child. For example, three tosses of the coin resulted in HHT, representing a family with two boys and one girl. Note that HTH and THH also represent two boys and one girl.

Suppose that when a prime number (P) is rolled on the number cube, it simulates a boy birth, and a non-prime (N) simulates a girl birth. Using such a number cube, list the outcomes that would simulate a boy birth, and those that simulate a girl birth. Are the boy and girl birth outcomes equally likely?

Suppose that one card is drawn from a regular deck of cards. A red card (R) simulates a boy birth, and a black card (B) simulates a girl birth. Describe how a family of three children could be simulated.

Example 2

Simulation provides an estimate for the probability that a family of three children would have three boys or three girls by performing three tosses of a fair coin many times. Each sequence of three tosses is called a trial. If a trial results in either HHH or TTT, then the trial represents all boys or all girls, which is the event that we are interested in. These trials would be called a *success*. If a trial results in any other order of H's and T's, then it is called a *failure*.

The estimate for the probability that a family has either three boys or three girls based on the simulation is the number of successes divided by the number of trials. Suppose 100 trials are performed, and that in those 100 trials, 28 resulted in either HHH or TTT. Then, the estimated probability that a family of three children has either three boys or three girls would be $\frac{28}{100} = 0.28$.

Exercises

Find an estimate of the probability that a family with three children will have exactly one girl using the following outcomes of 50 trials of tossing a fair coin three times per trial. Use H to represent a boy birth and T to represent a girl birth.

HHT	HTH	HHH	TTH	THT	THT	HTT	HHH	TTH	HHH
HHT	TTT	HHT	TTH	HHH	HTH	THH	TTT	THT	THT
THT	HHH	THH	HTT	HTH	TTT	HTT	HHH	TTH	THT
THH	HHT	TTT	TTH	HTT	THH	HTT	HTH	TTT	HHH
HTH	HTH	THT	TTH	TTT	HHT	HHT	THT	TTT	HTT

Example 3: Basketball Player

Suppose that, on average, a basketball player makes about three out of every four foul shots. In other words, she has a 75% chance of making each foul shot she takes. Since a coin toss produces equally likely outcomes, it could not be used in a simulation for this problem.

Instead, a number cube could be used by specifying that the numbers 1, 2, or 3 represent a hit, the number 4 represents a miss, and the numbers 5 and 6 would be ignored. Based on the following 50 trials of rolling a fair number cube, find an estimate of the probability that she makes five or six of the six foul shots she takes.

441323	342124	442123	422313	441243
124144	333434	243122	232323	224341
121411	321341	111422	114232	414411
344221	222442	343123	122111	322131
131224	213344	321241	311214	241131
143143	243224	323443	324243	214322
214411	423221	311423	142141	411312
343214	123131	242124	141132	343122
121142	321442	121423	443431	214433
331113	311313	211411	433434	323314

Lesson 11: Conducting a Simulation to Estimate the Probability of an Event

Classwork

Examples 1–2: Simulation

In the last lesson, we used coins, number cubes, and cards to carry out simulations. Another option is putting identical pieces of paper or colored disks into a container, mixing them thoroughly, and then choosing one.

For example, if a basketball player typically makes five out of eight foul shots, then a colored disk could be used to simulate a foul shot. A green disk could represent a made shot, and a red disk could represent a miss. You could put five green and three red disks in a container, mix them, and then choose one to represent a foul shot. If the color of the disk is green, then the shot is made. If the color of the disk is red, then the shot is missed. This procedure simulates one foul shot.

Using colored disks, describe how one at-bat could be simulated for a baseball player who has a batting average of 0.300. Note that a batting average of 0.300 means the player gets a hit (on average) three times out of every ten times at bat. Be sure to state clearly what a color represents.

Using colored disks, describe how one at-bat could be simulated for a player who has a batting average of 0.273. Note that a batting average of 0.273 means that on average, the player gets 273 hits out of $1{,}000$ at-bats.

Example 3: Using Random Number Tables

Why is using colored disks not practical for the situation described in Example 2? Another way to carry out a simulation is to use a random number table, or a random number generator. In a random number table, the digits $0, 1, 2, 3, 4, 5, 6, 7, 8$, and 9 occur equally often in the long run. Pages and pages of random numbers can be found online.

For example, here are three lines of random numbers. The space after every five digits is only for ease of reading. Ignore the spaces when using the table.

25256	65205	72597	00562	12683	90674	78923	96568	32177	33855
76635	92290	88864	72794	14333	79019	05943	77510	74051	87238
07895	86481	94036	12749	24005	80718	13144	66934	54730	77140

To use the random number table to simulate an at-bat for the 0.273 hitter in Exercise 2, you could use a three-digit number to represent one at bat. The three-digit numbers from 000–272 could represent a hit, and the three-digit numbers from 273–999 could represent a non-hit. Using the random numbers above and starting at the beginning of the first line, the first three-digit random number is 252, which is between 000 and 272, so that simulated at-bat is a hit. The next three-digit random number is 566, which is a non-hit.

Continuing on the first line of the random numbers above, what would the hit/non-hit outcomes be for the next six at-bats? Be sure to state the random number and whether it simulates a hit or non-hit.

Example 4: Baseball Player

A batter typically gets to bat four times in a ballgame. Consider the 0.273 hitter from the previous example. Use the following steps (and the random numbers shown above) to estimate that player's probability of getting at least three hits (three or four) in four times at-bat.

Describe what one trial is for this problem.

Describe when a trial is called a success and when it is called a failure.

Simulate 12 trials. (Continue to work as a class, or let students work with a partner.)

Use the results of the simulation to estimate the probability that a 0.273 hitter gets three or four hits in four times at-bat. Compare your estimate with other groups.

Example 5: Birth Month

In a group of more than 12 people, is it likely that at least two people, maybe more, will have the same birth month? Why? Try it in your class.

Now, suppose that the same question is asked for a group of only seven people. Are you likely to find some groups of seven people in which there is a match but other groups in which all seven people have different birth months? In the following exercise, you will estimate the probability that at least two people in a group of seven were born in the same month.

Exercises

What might be a good way to generate outcomes for the birth month problem—using coins, number cubes, cards, spinners, colored disks, or random numbers?

How would you simulate one trial of seven birth months?

How is a success determined for your simulation?

How is the simulated estimate determined for the probability that a least two in a group of seven people were born in the same month?

Lesson 12: Applying Probability to Make Informed Decisions

Classwork

Example 1: Number Cube

Your teacher gives you a number cube with numbers 1–6 on its faces. You have never seen that particular cube before. You are asked to state a theoretical probability model for rolling it once. A probability model consists of the list of possible outcomes (the sample space) and the theoretical probabilities associated with each of the outcomes. You say that the probability model might assign a probability of $\frac{1}{6}$ to each of the possible outcomes, but because you have never seen this particular cube before, you would like to roll it a few times. (Maybe it is a trick cube.) Suppose your teacher allows you to roll it 500 times, and you get the following results:

Outcome	1	2	3	4	5	6
Frequency	77	92	75	90	76	90

Exercises 1–2

If the equally likely model were correct, about how many of each outcome would you expect to see if the cube is rolled 500 times?

Based on the data from the 500 rolls, how often were odd numbers observed? How often were even numbers observed?

Example 2: Probability Model

Two black balls and two white balls are put in a small cup whose bottom allows the four balls to fit snugly. After shaking the cup well, two patterns of colors are possible as shown. The pattern on the left shows the similar colors are opposite each other, and the pattern on the right shows the similar colors are next to or adjacent to each other.

Philippe is asked to specify a probability model for the chance experiment of shaking the cup and observing the pattern. He thinks that because there are two outcomes—like heads and tails on a coin—that the outcomes should be equally likely. Sylvia isn't so sure that the equally likely model is correct, so she would like to collect some data before deciding on a model.

Exercise 3

Collect data for Sylvia. Carry out the experiment of shaking a cup that contains four balls, two black and two white, observing, and recording whether the pattern is opposite or adjacent. Repeat this process 20 times. Then, combine the data with that collected by your classmates.

Do your results agree with Philippe's equally likely model, or do they indicate that Sylvia had the right idea? Explain.

Exercises 4–5

There are three popular brands of mixed nuts. Your teacher loves cashews, and in his experience of having purchased these brands, he suggests that not all brands have the same percentage of cashews. One has around 20% cashews, one has 25%, and one has 35%.

Your teacher has bags labeled A, B, and C representing the three brands. The bags contain red beads representing cashews and brown beads representing other types of nuts. One bag contains 20% red beads, another 25% red beads, and the third has 35% red beads. You are to determine which bag contains which percentage of cashews. You cannot just open the bags and count the beads.

Work as a class to design a simulation. You need to agree on what an outcome is, what a trial is, what a success is, and how to calculate the estimated probability of getting a cashew. Base your estimate on 50 trials.

Your teacher will give your group one of the bags labeled A, B, or C. Using your plan from part (a), collect your data. Do you think you have the 20%, 25%, or 35% cashews bag? Explain.

Exercises 6–8

Suppose you have two bags, A and B, in which there are an equal number of slips of paper. Positive numbers are written on the slips. The numbers are not known, but they are whole numbers between 1 and 75, inclusive. The same number may occur on more than one slip of paper in a bag.

These bags are used to play a game. In this game, you choose one of the bags, and then choose one slip from that bag. If you choose bag A, and the number you choose from it is a prime number, then you win. If you choose bag B, and the number you choose from it is a power of 2, you win. Which bag should you choose?

Emma suggests that it doesn't matter which bag you choose because you don't know anything about what numbers are inside the bags. So she thinks that you are equally likely to win with either bag. Do you agree with her? Explain.

Aamir suggests that he would like to collect some data from both bags before making a decision about whether or not the model is equally likely. Help Aamir by drawing 50 slips from each bag, being sure to replace each one before choosing again. Each time you draw a slip, record whether it would have been a winner or not. Using the results, what is your estimate for the probability of drawing a prime number from bag A and drawing a power of 2 from bag B?

If you were to play this game, which bag would you choose? Explain why you would pick this bag.

Lesson 13: Populations, Samples, and Generalizing from a Sample to a Population

Classwork

In this lesson, you will learn about collecting data from a sample that is selected from a population. You will also learn about summary values for both a population and a sample and think about what can be learned about the population by looking at a sample from that population.

Exercises 1–4: Collecting Data

Describe what you would do if you had to collect data to investigate the following statistical questions using either a sample statistic or a population characteristic. Explain your reasoning in each case.

How might you collect data to answer the question, "Does the soup taste good?"

How might you collect data to answer the question, "How many movies do students in your class see in a month?"

How might you collect data to answer the question, "What is the median price of a home in our town?"

How might you collect data to answer the question, "How many pets do people own in my neighborhood?"

How might you collect data to answer the question, "What is the typical number of absences in math classes at your school on a given day?"

How might you collect data to answer the question, "What is the typical life span of a particular brand of flashlight battery?"

How might you collect data to answer the question, "What percentage of girls and of boys in your school have a curfew?"

How might you collect data to answer the question, "What is the most common blood type of students in my class?"

A *population* is the entire set of objects (people, animals, plants, etc.) from which data might be collected. A *sample* is a subset of the population. Numerical summary values calculated using data from an entire population are called *population characteristics*. Numerical summary values calculated using data from a sample are called *statistics*.

For which of the scenarios in Exercise 1 did you describe collecting data from a population and which from a sample?

Think about collecting data in the scenarios above. Give at least two reasons you might want to collect data from a sample rather than from the entire population.

Make up a result you might get in response to the situations in Exercise 1, and identify whether the result would be based on a population characteristic or a sample statistic.

Does the soup taste good?

How many movies do your classmates see in a month?

What is the median price of a home in our town?

How many pets do people in my neighborhood own?

What is the typical number of absences in math classes at your school on a given day?

What is the typical life span of a particular brand of flashlight batteries?

What percentage of girls and of boys in your school that have a curfew?

What is the most common blood type of my classmates?

Exercise 5: Population or Sample?

Indicate whether the following statements are summarizing data collected to answer a statistical question from a population or from a sample. Identify references in the statement as population characteristics or sample statistics.

54% of the responders to a poll at a university indicated that wealth needed to be distributed more evenly among people.

Are students in the Bay Shore school district proficient on the state assessments in mathematics? In 2013, after all the tests taken by the students in the Bay Shore schools were evaluated, over 52% of those students were at or above proficient on the state assessment.

Does talking on mobile phones while driving distract people? Researchers measured the reaction times of 38 study participants as they talked on mobile phones and found that the average level of distraction from their driving was rated 2.25 out of 5.

Did most people living in New York in 2010 have at least a high school education? Based on the data collected from all New York residents in 2010 by the United States Census Bureau, 84.6% of people living in New York had at least a high school education.

Were there more deaths than births in the United States between July 2011 and July 2012? Data from a health service agency indicated that there were 2% more deaths than births in the U.S. during that timeframe.

What is the fifth best-selling book in the United States? Based on the sales of books in the United States, the fifth best-selling book was *Oh, the Places You'll Go!* by Dr. Seuss.

Exercises 6–8: A Census

When data are collected from an entire population, it is called a census. The United States takes a census of its population every ten years, with the most recent one occurring in 2010. Go to http://ri.essortment.com/unitedstatesce_rlta.htm to find the history of the U.S. census.

Identify three things that you found to be interesting.

Why is the census important in the United States?

Go to the site: www.census.gov/2010census/popmap/ipmtext.php?fl=36.

Select the state of New York.

How many people were living in New York for the 2010 census?

Estimate the ratio of those 65 and older to those under 18 years old. Why is this important to think about?

Is the ratio a population characteristic or a statistic? Explain your thinking.

The American Community Survey (ACS) takes samples from a small percentage of the U.S. population in years between the censuses. (www.census.gov/acs/www/about_the_survey/american_community_survey/)

What is the difference between the way the ACS collects information about the U.S. population and the way the U.S. Census Bureau collects information?

In 2011, the ACS sampled workers living in New York about commuting to work each day. Why do you think these data are important for the state to know?

Suppose that from a sample of $200{,}000$ New York workers, $32{,}400$ reported traveling more than an hour to work each day. From this information, statisticians determined that between 16% and 16.4% of the workers in the state traveled more than an hour to work every day in 2011. If there were $8{,}437{,}512$ workers in the entire population, about how many traveled more than an hour to work each day?

Reasoning from a sample to the population is called making an inference about a population characteristic. Identify the statistic involved in making the inference in part (c).

The data about traveling time to work suggest that across the United States typically between 79.8% and 80% of commuters travel alone, 10% to 10.2% carpool, and 4.9% to 5.1% use public transportation. Survey your classmates to find out how a worker in their family gets to work. How do the results compare to the national data? What might explain any differences?

Lesson 14: Selecting a Sample

Classwork

As you learned in Lesson 13, sampling is a central concept in statistics. Examining every element in a population is usually impossible. So, research and articles in the media typically refer to a "sample" from a population. In this lesson, you will begin to think about how to choose a sample.

Exercises 1–2: What is Random?

1. Write down a sequence of heads/tails you think would typically occur if you tossed a coin 20 times. Compare your sequence to the ones written by some of your classmates. How are they alike? How are they different?

2. Working with a partner, toss a coin 20 times, and write down the sequence of heads and tails you get.
   a. Compare your results with your classmates.

   b. How are your results from actually tossing the coin different from the sequences you and your classmates wrote down?

   c. Toni claimed she could make up a set of numbers that would be random. What would you say to her?

\1 Casey at the Bat

The Outlook wasn't brilliant for the Mudville nine that day: The score stood four to two, **\2** with but one inning more to play. And then when Cooney died at first, and Barrows did the same, A **\3** sickly silence fell upon the patrons of the game.

A straggling few got up to go in deep despair. The **\4** rest Clung to that hope which springs eternal in the human breast; They thought, if only Casey could get but **\5** a whack at that—We'd put up even money, now, with Casey at the bat.

But Flynn preceded Casey, as **\6** did also Jimmy Blake, And the former was a lulu and the latter was a cake; So upon that stricken **\7** multitude grim melancholy sat, For there seemed but little chance of Casey's getting to the bat.

But Flynn let drive **\8** a single, to the wonderment of all, And Blake, the much despised, tore the cover off the ball; And when **\9** the dust had lifted, and the men saw what had occurred, There was Jimmy safe at second and Flynn a **\10** hugging third.

Then from five thousand throats and more there rose a lusty yell; It rumbled through the valley, it **\11** rattled in the dell; It knocked upon the mountain and recoiled upon the flat, For Casey, mighty Casey, was advancing **\12** to the bat**.**

There was ease in Casey's manner as he stepped into his place; There was pride in Casey's **\13** bearing and a smile on Casey's face. And when, responding to the cheers, he lightly doffed his hat, No stranger **\14** in the crowd could doubt 'twas Casey at the bat.

Ten thousand eyes were on him as he rubbed his **\15** hands with dirt**;** Five thousand tongues applauded when he wiped them on his shirt. Then while the writhing pitcher ground **\16** the ball into his hip, Defiance gleamed in Casey's eye, a sneer curled Casey's lip.

And now the leather covered **\17** sphere came hurtling through the air, And Casey stood a-watching it in haughty grandeur there. Close by the sturdy batsman **\18** the ball unheeded sped—"That ain't my style," said Casey. "Strike one," the umpire said.

From the benches, black with **\19** people, there went up a muffled roar, Like the beating of the storm waves on a stern and distant shore. **\20** "Kill him! Kill the umpire!" shouted someone on the stand; And it's likely they'd a-killed him had not Casey raised **\21** his hand.

With a smile of Christian charity great Casey's visage shone; He stilled the rising tumult; he bade the **\22** game go on; He signaled to the pitcher, and once more the spheroid flew; But Casey still ignored it, and **\23** the umpire said, "Strike two."

"Fraud!" cried the maddened thousands, and echo answered fraud; But one scornful look from Casey **\24** and the audience was awed. They saw his face grow stern and cold, they saw his muscles strain, And they **\25** knew that Casey wouldn't let that ball go by again.

The sneer is gone from Casey's lip, his teeth are **\26** clenched in hate; He pounds with cruel violence his bat upon the plate. And now the pitcher holds the ball, **\27** and now he lets it go, And now the air is shattered by the force of Casey's blow.

Oh, somewhere **\28** in this favored land the sun is shining bright; The band is playing somewhere, and somewhere hearts are light, And **\29** somewhere men are laughing, and somewhere children shout; But there is no joy in Mudville—mighty Casey has struck out.

by Ernest Lawrence Thayer

Exercises 3–11: Length of Words in the Poem *Casey at the Bat*

3. Suppose you wanted to learn about the lengths of the words in the poem *Casey at the Bat*. You plan to select a sample of eight words from the poem and use these words to answer the following statistical question: On average, how long is a word in the poem? What is the population of interest here?

4. Look at the poem, *Casey at the Bat,* by Ernest Thayer, and select eight words you think are representative of words in the poem. Record the number of letters in each word you selected. Find the mean number of letters in the words you chose.

5. A random sample is a sample in which every possible sample of the same size has an equal chance of being chosen. Do you think the set of words you wrote down was random? Why or why not?

\

6. Working with a partner, follow your teacher's instruction for randomly choosing eight words. Begin with the title of the poem, and count a hyphenated word as one word.
   a. Record the eight words you randomly selected, and find the mean number of letters in those words.

   b. Compare the mean of your random sample to the mean you found in Exercise 4. Explain how you found the mean for each sample.

7. As a class, compare the means from Exercise 4 and the means from Exercise 6. Your teacher will provide a chart to compare the means. Record your mean from Exercise 4 and your mean for Exercise 6 on this chart.

8. Do you think the means from Exercise 4 or the means from Exercise 6 are more representative of the mean of all of the words in the poem? Explain your choice.

9. The actual mean of the words in the poem *Casey at the Bat* is 4.2 letters. Based on the fact that the population mean is 4.2 letters, are the means from Exercise 4 or means from Exercise 6 a better representation of the mean of the population. Explain your answer.

10. How did population mean of 4.2 letters compare to the mean of your random sample from Exercise 6 and to the mean you found in Exercise 4?

11. Summarize how you would estimate the mean number of letters in the words of another poem based on what you learned in the above exercises.

Lesson 17: Sampling Variability

Classwork

Example 1: Estimating a Population Mean

The owners of a gym have been keeping track of how long each person spends at the gym. Eight hundred of these times (in minutes) are shown in the population tables located at the end of the lesson. These 800 times will form the *population* that you will investigate in this lesson.

Look at the values in the population. Can you find the longest time spent in the gym in the population? Can you find the shortest?

On average, roughly how long do you think people spend at the gym? In other words, by just looking at the numbers in the two tables, make an estimate of the population mean.

You could find the population mean by typing all 800 numbers into a calculator or a computer, adding them up, and dividing by 800. This would be extremely time-consuming, and usually it is not possible to measure every value in a population.

Instead of doing a calculation using every value in the population, we will use a *random sample* to find the mean of the sample. The sample mean will then be used as an estimate of the population mean.

Example 2: Selecting a Sample Using a Table of Random Digits

The table of random digits provided with this lesson will be used to select items from a population to produce a random sample from the population. The list of digits is determined by a computer program that simulates a random selection of the digits 0, 1, 2, 3, 4, 5, 6, 7, 8, or 9. Imagine that each of these digits is written on a slip of paper and placed in a bag. After thoroughly mixing the bag, one slip is drawn, and its digit is recorded in this list of random digits. The slip is then returned to the bag, and another slip is selected. The digit on this slip is recorded and then returned to the bag. The process is repeated over and over. The resulting list of digits is called a random number table.

How could you use a table of random digits to take a random sample?

Step 1: Place the table of random digits in front of you. Without looking at the page, place the eraser end of your pencil somewhere on the table. Start using the table of random digits at the number closest to where your eraser touched the paper. This digit and the following two specify which observation from the population tables will be the first observation in your sample.

For example, suppose the eraser end of your pencil lands on the twelfth number in row 3 of the random digit table. This number is 5, and the two following numbers are 1 and 4. This means that the first observation in your sample is observation number 514 from the population. Find observation number 514 in the population table. Do this by going to Row 51 and moving across to the column heading "4." This observation is 53, so the first observation in your sample is 53.

If the number from the random number table is any number 800 or greater, you will ignore this number and use the next three digits in the table.

Step 2: Continue using the table of random digits from the point you reached, and select the other four observations in your sample like you did above.

For example, continuing on from the position in the example given in Step 1,

- The next number from the random digit table is 716, and observation 716 is 63
- The next number from the random digit table is 565, and observation 565 is 31.
- The next number from the random digit table is 911, and there is no observation 911. So, we ignore these three digits.
- The next number from the random digit table is 928, and there is no observation 928. So, we ignore these three digits.
- The next number from the random digit table is 303, and observation 303 is 70.
- The next number from the random digit table is 677, and observation 677 is 42.

Exercises 1–4

Initially, you will select just five values from the population to form your sample. This is a very small sample size, but it is a good place to start to understand the ideas of this lesson.

1. Use the table of random numbers to select five values from the population of times. What are the five observations in your sample?

2. For the sample that you selected, calculate the sample mean.

3. You selected a random sample and calculated the sample mean in order to estimate the population mean. Do you think that the mean of these five observations is exactly correct for the population mean? Could the population mean be greater than the number you calculated? Could the population mean be less than the number you calculated?

4. In practice, you only take one sample in order to estimate a population characteristic. But, for the purposes of this lesson, suppose you were to take another random sample from the same population of

times at the gym. Could the new sample mean be closer to the population mean than the mean of these five observations? Could it be farther from the population mean?

Exercises 5–7

As a class, you will now investigate sampling variability by taking several samples from the same population. Each sample will have a different sample mean. This variation provides an example of sampling variability.

5. Place the table of random digits in front of you, and without looking at the page, place the eraser end of your pencil somewhere on the table of random numbers. Start using the table of random digits at the number closest to where your eraser touches the paper. This digit and the following two specify which observation from the population tables will be the first observation in your sample. Write this three-digit number and the corresponding data value from the population in the space below.

6. Continue moving to the right in the table of random digits from the place you ended in Exercise 5. Use three digits at a time. Each set of three digits specifies which observation in the population is the next number in your sample. Continue until you have four more observations, and write these four values in the space below.

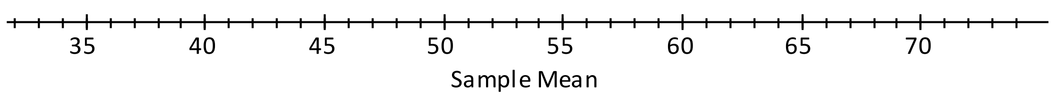
7. Calculate the mean of the five values that form your sample. Round your answer to the nearest tenth. Show your work and your sample mean in the space below.

Exercises 8–11

You will now use the sample means from Exercise 7 from the entire class to make a dot plot.

8. Write the sample means for everyone in the class in the space below.

9. Use all the sample means to make a dot plot using the axis given below. (Remember, if you have repeated or close values, stack the dots one above the other.)



10. What do you see in the dot plot that demonstrates sampling variability?

11. Remember that in practice you only take one sample. (In this lesson, many samples were taken in order to demonstrate the concept of sampling variability.) Suppose that a statistician plans to take a random sample of size 5 from the population of times spent at the gym and that he will use the sample mean as an estimate of the population mean. Approximately how far can the statistician expect the sample mean to be from the population mean?

Population

	0	1	2	3	4	5	6	7	8	9
00	45	58	49	78	59	36	52	39	70	51
01	50	45	45	66	71	55	65	33	60	51
02	53	83	40	51	83	57	75	38	43	77
03	49	49	81	57	42	36	22	66	68	52
04	60	67	43	60	55	63	56	44	50	58
05	64	41	67	73	55	69	63	46	50	65
06	54	58	53	55	51	74	53	55	64	16
07	28	48	62	24	82	51	64	45	41	47
08	70	50	38	16	39	83	62	50	37	58
09	79	62	45	48	42	51	67	68	56	78
10	61	56	71	55	57	77	48	65	61	62
11	65	40	56	47	44	51	38	68	64	40
12	53	22	73	62	82	78	84	50	43	43
13	81	42	72	49	55	65	41	92	50	60
14	56	44	40	70	52	47	30	9	58	53
15	84	64	64	34	37	69	57	75	62	67
16	45	58	49	78	59	36	52	39	70	51
17	50	45	45	66	71	55	65	33	60	51
18	53	83	40	51	83	57	75	38	43	77
19	49	49	81	57	42	36	22	66	68	52
20	60	67	43	60	55	63	56	44	50	58
21	64	41	67	73	55	69	63	46	50	65
22	54	58	53	55	51	74	53	55	64	16
23	28	48	62	24	82	51	64	45	41	47
24	70	50	38	16	39	83	62	50	37	58
25	79	62	45	48	42	51	67	68	56	78
26	61	56	71	55	57	77	48	65	61	62
27	65	40	56	47	44	51	38	68	64	40
28	53	22	73	62	82	78	84	50	43	43
29	81	42	72	49	55	65	41	92	50	60
30	56	44	40	70	52	47	30	9	58	53
31	84	64	64	34	37	69	57	75	62	67
32	45	58	49	78	59	36	52	39	70	51
33	50	45	45	66	71	55	65	33	60	51
34	53	83	40	51	83	57	75	38	43	77
35	49	49	81	57	42	36	22	66	68	52
36	60	67	43	60	55	63	56	44	50	58
37	64	41	67	73	55	69	63	46	50	65
38	54	58	53	55	51	74	53	55	64	16
39	28	48	62	24	82	51	64	45	41	47

Population, continued

	0	1	2	3	4	5	6	7	8	9
40	53	70	59	62	33	31	74	44	46	68
41	37	51	84	47	46	33	53	54	70	74
42	35	45	48	45	56	60	66	60	65	57
43	42	81	67	64	60	79	46	48	67	56
44	41	21	41	58	48	38	50	53	73	38
45	35	28	43	43	55	39	75	45	68	36
46	64	31	31	40	84	79	47	63	48	46
47	34	36	54	61	33	16	50	60	52	55
48	53	52	48	47	77	37	66	51	61	64
49	40	44	45	22	36	64	50	49	64	39
50	45	69	67	33	55	61	62	38	51	43
51	55	39	46	56	53	50	44	42	40	60
52	11	36	56	69	72	73	71	48	58	52
53	81	47	36	54	81	59	50	42	80	69
54	40	43	30	54	61	13	73	65	52	40
55	71	78	71	61	54	79	63	47	49	73
56	53	70	59	62	33	31	74	44	46	68
57	37	51	84	47	46	33	53	54	70	74
58	35	45	48	45	56	60	66	60	65	57
59	42	81	67	64	60	79	46	48	67	56
60	41	21	41	58	48	38	50	53	73	38
61	35	28	43	43	55	39	75	45	68	36
62	64	31	31	40	84	79	47	63	48	46
63	34	36	54	61	33	16	50	60	52	55
64	53	52	48	47	77	37	66	51	61	64
65	40	44	45	22	36	64	50	49	64	39
66	45	69	67	33	55	61	62	38	51	43
67	55	39	46	56	53	50	44	42	40	60
68	11	36	56	69	72	73	71	48	58	52
69	81	47	36	54	81	59	50	42	80	69
70	40	43	30	54	61	13	73	65	52	40
71	71	78	71	61	54	79	63	47	49	73
72	53	70	59	62	33	31	74	44	46	68
73	37	51	84	47	46	33	53	54	70	74
74	35	45	48	45	56	60	66	60	65	57
75	42	81	67	64	60	79	46	48	67	56
76	41	21	41	58	48	38	50	53	73	38
77	35	28	43	43	55	39	75	45	68	36
78	64	31	31	40	84	79	47	63	48	46
79	34	36	54	61	33	16	50	60	52	55

Table of Random Digits

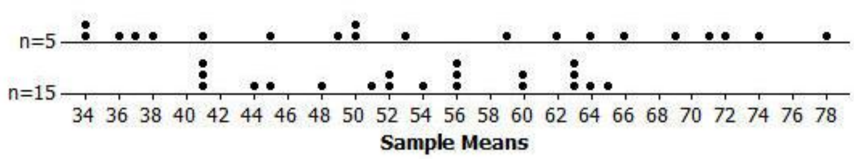
Row																				
1	6	6	7	2	8	0	0	8	4	0	0	4	6	0	3	2	2	4	6	8
2	8	0	3	1	1	1	1	2	7	0	1	9	1	2	7	1	3	3	5	3
3	5	3	5	7	3	6	3	1	7	2	5	5	1	4	7	1	6	5	6	5
4	9	1	1	9	2	8	3	0	3	6	7	7	4	7	5	9	8	1	8	3
5	9	0	2	9	9	7	4	6	3	6	6	3	7	4	2	7	0	0	1	9
6	8	1	4	6	4	6	8	2	8	9	5	5	2	9	6	2	5	3	0	3
7	4	1	1	9	7	0	7	2	9	0	9	7	0	4	6	2	3	1	0	9
8	9	9	2	7	1	3	2	9	0	3	9	0	7	5	6	7	1	7	8	7
9	3	4	2	2	9	1	9	0	7	8	1	6	2	5	3	9	0	9	1	0
10	2	7	3	9	5	9	9	3	2	9	3	9	1	9	0	5	5	1	4	2
11	0	2	5	4	0	8	1	7	0	7	1	3	0	4	3	0	6	4	4	4
12	8	6	0	5	4	8	8	2	7	7	0	1	0	1	7	1	3	5	3	4
13	4	2	6	4	5	2	4	2	6	1	7	5	6	6	4	0	8	4	1	2
14	4	4	9	8	7	3	4	3	8	2	9	1	5	3	5	9	8	9	2	9
15	6	4	8	0	0	0	4	2	3	8	1	8	4	0	9	5	0	9	0	4
16	3	2	3	8	4	8	8	6	2	9	1	0	1	9	9	3	0	7	3	5
17	6	6	7	2	8	0	0	8	4	0	0	4	6	0	3	2	2	4	6	8
18	8	0	3	1	1	1	1	2	7	0	1	9	1	2	7	1	3	3	5	3
19	5	3	5	7	3	6	3	1	7	2	5	5	1	4	7	1	6	5	6	5
20	9	1	1	9	2	8	3	0	3	6	7	7	4	7	5	9	8	1	8	3
21	9	0	2	9	9	7	4	6	3	6	6	3	7	4	2	7	0	0	1	9
22	8	1	4	6	4	6	8	2	8	9	5	5	2	9	6	2	5	3	0	3
23	4	1	1	9	7	0	7	2	9	0	9	7	0	4	6	2	3	1	0	9
24	9	9	2	7	1	3	2	9	0	3	9	0	7	5	6	7	1	7	8	7
25	3	4	2	2	9	1	9	0	7	8	1	6	2	5	3	9	0	9	1	0
26	2	7	3	9	5	9	9	3	2	9	3	9	1	9	0	5	5	1	4	2
27	0	2	5	4	0	8	1	7	0	7	1	3	0	4	3	0	6	4	4	4
28	8	6	0	5	4	8	8	2	7	7	0	1	0	1	7	1	3	5	3	4
29	4	2	6	4	5	2	4	2	6	1	7	5	6	6	4	0	8	4	1	2
30	4	4	9	8	7	3	4	3	8	2	9	1	5	3	5	9	8	9	2	9
31	6	4	8	0	0	0	4	2	3	8	1	8	4	0	9	5	0	9	0	4
32	3	2	3	8	4	8	8	6	2	9	1	0	1	9	9	3	0	7	3	5
33	6	6	7	2	8	0	0	8	4	0	0	4	6	0	3	2	2	4	6	8
34	8	0	3	1	1	1	1	2	7	0	1	9	1	2	7	1	3	3	5	3
35	5	3	5	7	3	6	3	1	7	2	5	5	1	4	7	1	6	5	6	5
36	9	1	1	9	2	8	3	0	3	6	7	7	4	7	5	9	8	1	8	3
37	9	0	2	9	9	7	4	6	3	6	6	3	7	4	2	7	0	0	1	9
38	8	1	4	6	4	6	8	2	8	9	5	5	2	9	6	2	5	3	0	3
39	4	1	1	9	7	0	7	2	9	0	9	7	0	4	6	2	3	1	0	9
40	9	9	2	7	1	3	2	9	0	3	9	0	7	5	6	7	1	7	8	7

Lesson 18: Sampling Variability and the Effect of Sample Size

Classwork

Example 1: Sampling Variability

The previous lesson investigated the statistical question, “What is the typical time spent at the gym?” by selecting random samples from the population of 800 gym members. Two different dot plots of sample means calculated from random samples from the population are displayed below. The first dot plot represents the means of 20 samples with each sample having 5 data points. The second dot plot represents the means of 20 samples with each sample having 15 data points.



Based on the first dot plot, Jill answered the statistical question by indicating the mean time people spent at the gym was between 34 and 78 minutes. She decided that a time approximately in the middle of that interval would be her estimate of the mean time the 800 people spent at the gym. She estimated 52 minutes. Scott answered the question using the second dot plot. He indicated that the mean time people spent at the gym was between 41 and 65 minutes. He also selected a time of 52 minutes to answer the question.

- Describe the differences in the two dot plots.

- Which dot plot do you feel more confident in using to answer the statistical question? Explain your answer.

- In general, do you want sampling variability to be large or small? Explain.

Exercises 1–3

In the previous lesson, you saw a population of 800 times spent at the gym. You will now select a random sample of size 15 from that population. You will then calculate the sample mean.

1. Start by selecting a three-digit number from the table of random digits. Place the random digit table in front of you. Without looking at the page, place the eraser end of your pencil somewhere on the table of random digits. Start using the table of random digits at the digit closest to your eraser. This digit and the following two specify which observation from the population will be the first observation in your sample. Write the value of this observation in the space below. (Discard any three-digit number that is 800 or larger, and use the next three digits from the random digit table.)

2. Continue moving to the right in the table of random digits from the point that you reached in Exercise 1. Each three-digit number specifies a value to be selected from the population. Continue in this way until you have selected 14 more values from the population. This will make 15 values altogether. Write the values of all 15 observations in the space below.

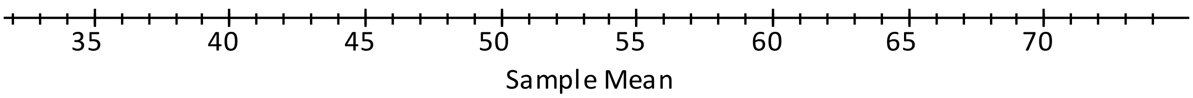
3. Calculate the mean of your 15 sample values. Write the value of your sample mean below. Round your answer to the nearest tenth. (Be sure to show your work.)

Exercises 4–6

You will now use the sample means from Exercise 3 for the entire class to make a dot plot.

4. Write the sample means for everyone in the class in the space below.

5. Use all the sample means to make a dot plot using the axis given below. (Remember, if you have repeated values or values close to each other, stack the dots one above the other.)



6. In the previous lesson, you drew a dot plot of sample means for samples of size 5. How does the dot plot above (of sample means for samples of size 15) compare to the dot plot of sample means for samples of size 5? For which sample size (5 or 15) does the sample mean have the greater sampling variability?

This exercise illustrates the notion that the greater the sample size, the smaller the sampling variability of the sample mean.

Exercises 7–8

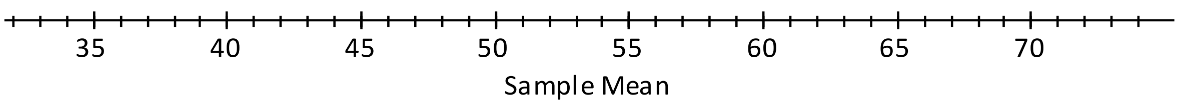
7. Remember that in practice you only take one sample. Suppose that a statistician plans to take a random sample of size 15 from the population of times spent at the gym and will use the sample mean as an estimate of the population mean. Based on the dot plot of sample means that your class collected from the population, approximately how far can the statistician expect the sample mean to be from the population mean? (The actual population mean is 53.9 minutes.)

8. How would your answer in Exercise 7 compare to the equivalent mean of the distances for a sample of size 5?

Exercises 9–11

Suppose everyone in your class selected a random sample of size 25 from the population of times spent at the gym.

9. What do you think the dot plot of the class's sample means would look like? Make a sketch using the axis below.



10. Suppose that a statistician plans to estimate the population mean using a sample of size 25. According to your sketch, approximately how far can the statistician expect the sample mean to be from the population mean?

11. Suppose you have a choice of using a sample of size 5, 15, or 25. Which of the three makes the sampling variability of the sample mean the smallest? Why would you choose the sample size that makes the sampling variability of the sample mean as small as possible?

Lesson 22: Using Sample Data to Compare the Means of Two or More Populations

Classwork

In previous lessons, you worked with one population. Many statistical questions involve comparing two populations. For example:

- On average, do boys and girls differ on quantitative reasoning?
- Do students learn basic arithmetic skills better with or without calculators?
- Which of two medications is more effective in treating migraine headaches?
- Does one type of car get better mileage per gallon of gasoline than another type?
- Does one type of fabric decay in landfills faster than another type?
- Do people with diabetes heal more slowly than people who do not have diabetes?

In this lesson, you will begin to explore how big of a difference there needs to be in sample means in order for the difference to be considered meaningful. The next lesson will extend that understanding to making informal inferences about population differences.

The **Mean Absolute Deviation (MAD)** of a set of data is the average distance between each data value and the mean.

The steps to find the MAD include:

1. find the mean (average)
2. find the difference between each data value and the mean
3. take the absolute value of each difference
4. find the mean (average) of these differences

Examples 1–3

Tamika's mathematics project is to see whether boys or girls are faster in solving a KenKen-type puzzle. She creates a puzzle and records the following times that it took to solve the puzzle (in seconds) for a random sample of 10 boys from her school and a random sample of 11 girls from her school:

													Mean	MAD
Boys	39	38	27	36	40	27	43	36	34	33			35.3	4.04
Girls	41	41	33	42	47	38	41	36	36	32	46		39.4	3.96

1. On the same scale, draw dot plots for the boys' data and for the girls' data. Comment on the amount of overlap between the two dot plots. How are the dot plots the same, and how are they different?

2. Compare the variability in the two data sets using the MAD (mean absolute deviation). Is the variability in each sample about the same? Interpret the MAD in the context of the problem.

3. In the previous lesson, you learned that a difference between two sample means is considered to be meaningful if the difference is more than what you would expect to see just based on sampling variability. The difference in the sample means of the boys' times and the girls' times is 4.1 seconds (39.4 seconds − 35.3 seconds). This difference is approximately 1 MAD.

   a. If 4 sec. is used to approximate the values of 1 MAD for both boys and for girls, what is the interval of times that are within 1 MAD of the sample mean for boys?

b. Of the 10 sample means for boys, how many of them are within that interval?

c. Of the 11 sample means for girls, how many of them are within the interval you calculated in part (a)?

d. Based on the dot plots, do you think that the difference between the two sample means is a meaningful difference? That is, are you convinced that the mean time for all girls at the school (not just this sample of girls) is different from the mean time for all boys at the school? Explain your choice based on the dot plots.

Examples 4–7

How good are you at estimating a minute? Work in pairs. Flip a coin to determine which person in the pair will go first. One of you puts your head down and raises your hand. When your partner says "start," keep your head down and your hand raised. When you think a minute is up, put your hand down. Your partner will record how much time has passed. Note that the room needs to be quiet. Switch roles except this time you talk with your partner during the period when the person with his head down is indicating when he thinks a minute is up. Note that the room will not be quiet.

Group	Estimate for a minute													
Quiet														
Talking														

Use your class data to complete the following.

4. Calculate the mean minute time for each group. Then, find the difference between the "quiet" mean and the "talking" mean.

5. On the same scale, draw dot plots of the two data distributions, and discuss the similarities and differences in the two distributions.

6. Calculate the mean absolute deviation (MAD) for each data set. Based on the MADs, compare the variability in each sample. Is the variability about the same? Interpret the MADs in the context of the problem.

7. Based on your calculations, is the difference in mean time estimates meaningful? Part of your reasoning should involve the number of MADs that separate the two sample means. Note that if the MADs differ, use the larger one in determining how many MADs separate the two means.

Lesson 23: Using Sample Data to Compare the Means of Two or More Populations

Classwork

In the previous lesson, you described how far apart the means of two data sets are in terms of the MAD (mean absolute deviation), a measure of variability. In this lesson, you will extend that idea to informally determine when two sample means computed from random samples are far enough apart from each other to imply that the population means also differ in a *meaningful* way. Recall that a *meaningful* difference between two means is a difference that is greater than would have been expected just due to sampling variability.

Example 1: Texting

With texting becoming so popular, Linda wanted to determine if middle school students memorize *real* words more or less easily than *fake* words. For example, real words are *food*, *car*, *study*, *swim*; whereas fake words are *stk*, *fonw*, *cqur*, *ttnsp*. She randomly selected 28 students from all middle school students in her district and gave half of them a list of 20 real words and the other half a list of 20 fake words.

a. How do you think Linda might have randomly selected 28 students from all middle school students in her district?

b. Why do you think Linda selected the students for her study randomly? Explain.

c. She gave the selected students one minute to memorize their lists after which they were to turn the lists over and, after two minutes, write down all the words that they could remember. Afterward, they calculated the number of correct words that they were able to write down. Do you think a penalty should be given for an incorrect word written down? Explain your reasoning.

Exercises 1–4

Suppose the data (number of correct words recalled) she collected were as follows

For students given the real words list: $8, 11, 12, 8, 4, 7, 9, 12, 12, 9, 14, 11, 5, 10$.

For students given the fake words list: $3, 5, 4, 4, 4, 7, 11, 9, 7, 7, 1, 3, 3, 7$.

1. On the same scale, draw dot plots for the two data sets.

2. From looking at the dot plots, write a few sentences comparing the distribution of the number of correctly recalled real words with the distribution of number of correctly recalled fake words. In particular, comment on which type of word, if either, that students recall better. Explain.

3. Linda made the following calculations for the two data sets:

	Mean	MAD
Real words recalled	9.43	2.29
Fake words recalled	5.36	2.27

In the previous lesson, you calculated the number of MADs that separated two sample means. You used the larger MAD to make this calculation if the two MADs were not the same. How many MADs separate the mean number of real words recalled and the mean number of fake words recalled for the students in the study?

4. In the last lesson, our work suggested that if the number of MADs that separate the two sample means is 2 or more, then it is reasonable to conclude that not only do the means differ in the samples, but that the means differ in the populations as well. If the number of MADs is less than 2, then you can conclude that the difference in the sample means might just be sampling variability and that there may not be a meaningful difference in the population means. Using these criteria, what can Linda conclude about the difference in population means based on the sample data that she collected? Be sure to express your conclusion in the context of this problem.

Example 2

Ken, an eighth-grade student, was interested in doing a statistics study involving sixth-grade and eleventh-grade students in his school district. He conducted a survey on four numerical variables and two categorical variables (grade level and gender). His Excel population database for the 265 sixth graders and 175 eleventh graders in his district has the following description:

Column	Name	Description
1	ID	ID numbers are from 1 through 440 1–128 Sixth-grade females 129–265 Sixth-grade males 266–363 Eleventh-grade females 364–440 Eleventh-grade males
2	Texting	Number of minutes per day text (whole number)
3	ReacTime	Time in seconds to respond to a computer screen stimulus (two decimal places)
4	Homework	Total number of hours per week spent on doing homework (one decimal place)
5	Sleep	Number of hours per night sleep (one decimal place)

a. Ken decides to base his study on a random sample of 20 sixth graders and a random sample of 20 eleventh graders. The sixth graders have IDs 1–265, and the eleventh graders are numbered 266–440. Advise him on how to randomly sample 20 sixth graders and 20 eleventh graders from his data file.

Suppose that from a random number generator:

The random ID numbers for his 20 sixth graders:
231 15 19 206 86 183 233 253 142 36 195 139 75 210 56 40 66 114 127 9

The random ID numbers for his 20 eleventh graders:
391 319 343 426 307 360 289 328 390 350 279 283 302 287 269 332 414 267 428

b. For each set, find the homework hours data from the population database that corresponds to these randomly selected ID numbers.

c. On the same scale, draw dot plots for the two sample data sets.

d. From looking at the dot plots, list some observations comparing the number of hours per week that sixth graders spend on doing homework and the number of hours per week that eleventh graders spend on doing homework.

e. Calculate the mean and MAD for each of the data sets. How many MADs separate the two sample means? (Use the larger MAD to make this calculation if the sample MADs are not the same.)

	Mean (hr.)	MAD (hr.)
Sixth Grade		
Eleventh Grade		

f. Ken recalled Linda suggesting that if the number of MADs is greater than or equal to 2, then it would be reasonable to think that the population of all sixth-grade students in his district and the population of all eleventh-grade students in his district have different means. What should Ken conclude based on his homework study?